Image Retrieval Using Cubic Splies Neural Networks

Samy Sadek [1], Ayoub Al-Hamadi [1], Bernd Michaelis [1], and Usama Sayed [2]
[1]Institute for Electronics, Signal Processing and Communications,
Otto-von-Guericke University Magdeburg, Germany
[2]Electrical Engineering Department,
Assiut University, Assiut, Egypt
{Samy.Bakheet, Ayoub.Al-Hamadi}@ovgu.de

***Abstract*-- Most of the approaches of Content-Based Image Retrieval (CBIR) presume a linear relationship between different image features, and the efficiency of such systems was limited due to the difficulty in representing high-level concepts using low-level features. In this paper, a new architecture for a CBIR system is proposed; the Splines Neural Network-based Image Retrieval (SNNIR) system. SNNIR makes use of a rapid and precise network model that employs a cubic-splines activation function. By using the cubic-splines network, the proposed system could determine nonlinear relationship between images features so that more accurate similarity comparison between images can be supported. Experimental results show that the proposed system achieves high accuracy and effectiveness in terms of precision and recall compared to other CBIR systems.**

***Index Term* — Splines neural network, Feature extraction, Content-based image retrieval.**

1. INTRODUCTION

While a variety of methods to improve the performance of Artificial Neural Network (ANN) have been investigated via optimizing training methods, learn parameters, or network structure, few works have been done towards using adaptive activation functions other than sigmoidal functions. It has been shown that a complexity of ANN, both structural, in terms of interconnections, and computational, in terms of the number of multiplications, is a bottleneck for many real-time applications. However, it is known that under certain regularity conditions, the ANN representation capabilities depend on the number of free parameters, whatever the structure of the network [1]. Therefore, adaptable activation functions can reduce the number of interconnections and consequently the overall network complexity, since they currently contain free parameters. In digital ANN implementation, the computational complexity of the forward phase can be further reduced when the activation function is implemented through a lookup table (LUT) [2]. The LUT access time is noticeably independent from the function shape.

The conventional neuron model, proposed by McCulloch and Pitts in 1943 [3], is composed of a linear combiner followed by a nonlinear function (activation function) with hard-limiting characteristics. Recently, in order to develop gradientbased learning algorithms, such hard-limiters are substituted with nonlinear differentiable functions [4]. It is well known that the network behavior, as that shown by the multilayer perceptron (MLP), greatly depends on the shape of these activation functions. Sigmoidal functions are the most common in ANN applications [5], [6], however, other kinds of functions, that dependent on some free parameters, are used to improve the NN's representation capabilities.

As a further step, some authors proposed a MLP with adaptive-slope sigmoidal activation function [7]. In [8], a more general approach put forward: an adaptive generalized hyperbolic tangent function with two free parameters, the slope and the saturation level (or gain), is suggested. Moreover, NN's universal approximation capabilities are guaranteed for a large class activation functions [9, 10] and the behavior of different classes of activation functions have been studied in depth by several authors [11,12,13,14].

The high-speed growth in the number of large-scale image repositories in many domains such as multimedia libraries, document archives, medical image management, biometrics, and environmental monitoring has bought the need for efficient CBIR mechanisms [15]. Loads of papers have been published in the last few years in this area. For instance, Wan and Kuo [16] have proposed an approach for image retrieval with hierarchical color clustering. In [17], a system that mimics the human brain for CBIR was presented. CBIR systems still do not carry out as well as their text counterparts [18]. Image retrieval systems have traditionally relied on annotations or captions associated with the images for indexing the retrieval system. The labor-intensive task of indexing and cataloging the images in these collections has conventionally been performed manually, a process that can be subjective and prone to errors.

The last decades have seen abundant advancements in the area of CBIR [15]. Although CBIR approaches have demonstrated success in moderately constrained domains including pathology, dermatology, chest radiology, and mammography, they have verified poor performance when applied to databases with a wide spectrum of imaging modalities, anatomies and pathologies [18, 19, 20, 21]. Image retrieval performance has shown comprehensible improvement by fusing the results of textual and visual techniques. This has particularly been shown to improve early precision [22, 23].

In this paper, a new neural system for image retrieval named

 IJENS

splines neural network image retrieval (SNNIR) is presented. The proposed system utilizes an adaptive neural network model called splines neural network (SNN). The splines neural network enables the system to determine nonlinear relationship between different features in images. Results of the proposed system show that it is more effective and efficient to retrieve visual-similar images for a set of images with same conception can be retrieved.

The rest of the paper is organized as follow. In section 2, the cubic-splines neural network architecture is presented. Color features are described in section 3. Section 4 investigates the architecture of SNNIR system. In section 5, experimental results are reported to show the performance of the proposed approach. Concluding remarks are offered in section 6.

2. CUBIC-SPLINES NEURAL NETWORKS

The activation function simulates the correlation between the action potential of the inputs and the output of the neuron. Artificial Neural Networks (ANN) implementations are frequently using well-known activation functions like the sigmoidal functions. In this paper we employ an adaptive activation function for the hidden neurons out of a pool of standard functions called cubic-splines function to increase flexibility. In general, choosing a function as activation function should take into account these aspects: Simple implementation, fast computation, and the used function should be partially refineable. These requirements lead to the mathematical field of interpolating polynomials, a part of numerical analysis. Cubic-splines function consists of third degree polynomials. Each two consecutive data points (i.e. control points) are connected by a specific polynomial. Mathematically, cubic-splines function is defined as:

$$S(x) = s_k(x) = \sum_{i=0}^{3} s_{k,i}(x - x_k)^i \quad \forall\, x \in [x_k, x_{k+1}],\, k = 1,2,\dots,n-1 \tag{1}$$

where $s_{k,i}$ are the coefficients of the cubic-splines function.

Now, a system with four equations has to be set up to determine the coefficients. The required system can be constructed by using the constraints of the cubic-splines functions (i.e. the first and second derivatives). By combining Eq. (1) and the splines' constraints we finally get the following matrix of equations (see [24] for more details).

$$\begin{pmatrix} s_{k,0} \\ s_{k,1} \\ s_{k,2} \\ s_{k,3} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\alpha & 0 & \alpha & 0 \\ 2\alpha & \alpha-3 & 3-2\alpha & -\alpha \\ -\alpha & 2-\alpha & \alpha-2 & \alpha \end{pmatrix} X \tag{2}$$

$$X = \begin{pmatrix} x_{k-1} \\ x_k \\ x_{k+1} \\ x_{k+2} \end{pmatrix}, k = 2,3,\dots,n-1.$$

In Eq. (1) above, α is the tension parameter of cubic-splines function that determines how sharply the curve bends at the control points. The common value for α is 0.5. Therefore, if we substitute $\alpha = 0.5$ into Eq. (2), we can determine the cubic-splines function using Eq. (1) as follows:

$$S(u) = \frac{1}{2}(1 \quad u \quad u^2 \quad u^3) A X \tag{3}$$

$$A = \begin{pmatrix} 0 & 2 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 2 & -5 & 4 & -1 \\ -1 & 3 & -3 & 1 \end{pmatrix}$$

$$0 \le u \le 1.$$

3. COLOR FEATURES

Image retrieval systems usually use some features that characterize the image. In the existing content-based image retrieval systems the most common features are color, shape, and texture. The distributions of colors of images are commonly used characteristics in tasks like content-based image retrieval. The color distribution can be presented as a color histogram.

Color histograms are frequently used in image classification and retrieval systems, in [25], [26], for example, the original idea is based on using color histograms in indexing of large image databases. After that, color histograms have also been used in recognition of objects in the images [27]. In this work, feature vectors depend on the following color descriptors: color histograms, color coherence and color moments, as described below.

3.1. HSV histogram

RGB histogram is still not a satisfactory choice for image retrieval system. Therefore, another color space is necessary. HSV color space is one of the good alternatives.

The Hue/Saturation/Value model was created by A. R. Smith in 1978. It is based on such intuitive color characteristics as tint, shade and tone. HSV offers improved perceptual uniformity. It represents with equal emphasis the three color variants that characterize color. This separation is attractive because color image processing performed independently on the color channels does not introduce false colors. HSV components extracted from RGB depend on which of the RGB components is largest and which is smallest, and are briefly given by C++ pseudo-code below:

```
min = MIN( r, g, b );
max = MAX( r, g, b );
v = max;
delta = max - min;
if( max != 0)
    s = delta /max;
    if( r == max ) h = (g - b)/delta;
    else if( g == max ) h = 2 + (b-r)/delta;
    else h = 4 + (r - g) / delta;
    h = h*60;
    if( h < 0 ) h = h+360;
else
    s = 0;
    h = -1;
```

3.2. Color coherence

Color Coherence (CC) is a different way of incorporating spatial information into the color histogram. Each histogram bin is partitioned into two types: coherent, if it belongs to a large uniformly-colored region, or incoherent, if it does not. Let α_i, β_i denote the number of coherent and incoherent pixels in the *i-th* color bin respectively. Then, the CC vector is defined as:

$$\langle(\alpha_1, \beta_1), (\alpha_2, \beta_2), \ldots, (\alpha_n, \beta_n)\rangle \quad (4)$$

3.3. Color moments

The origin of color moments lays in the assumption that the distribution of color in an image can be interpreted as a probability distribution. Probability distributions are characterized by a number of unique moments (e.g. Normal distributions are differentiated by their mean and variance). Hence, if the color in an image follows a certain probability distribution, the moments of that distribution can then be used as features to identify that image based on color. The three central moments of an image are mean, standard deviation and skewness. Mathematically, the three color moments can be defined as:

$$\mu_k = \frac{1}{n}\sum_{i=1}^{n} p_i^k \quad (5.a)$$

$$\sigma_k = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(p_i^k - \mu_k)^2} \quad (5.b)$$

$$s_k = \sqrt[3]{\frac{1}{n}\sum_{i=1}^{n}(p_i^k - \mu_k)^3} \quad (5.c)$$

where p_i^k is the value of the *k-th* color channel at the *i-th* pixel, and *n* is the size of image.

4. SPLINES-NN IMAGE RETRIEVAL SYSTEM

In this section, the structure of the proposed SNNIR system is described. Fig. 1 shows the main components of the proposed system. There are two stages of the control flow. The first deals with learning the splines NN and saving feature vectors, while the second deals with the query image. During the learning stage, a set of images in the database has been grouped into predefined categories.

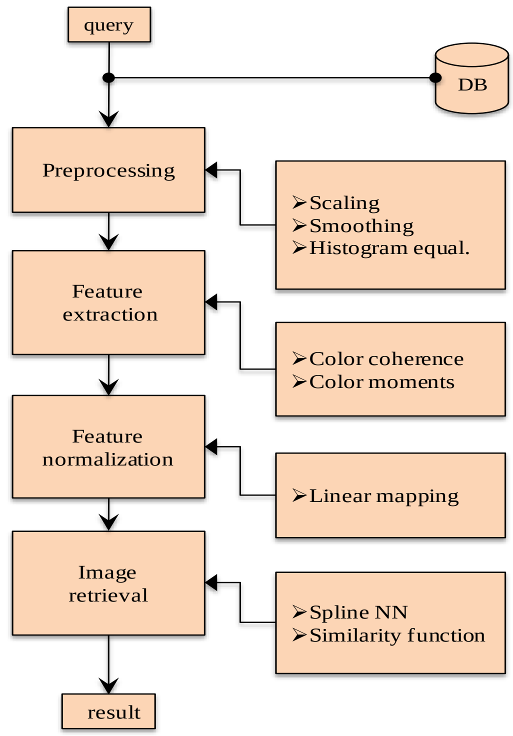

Fig. 1. Block diagram for the proposed SNNIR system

The main components of the SNNIR system shown above in Fig. 1 are described as follows.

4.1. Pre-processing

A preliminary step of creating a symbolic representation of the source images is required before applying any data retrieval methods. The images are thus normalized by bringing them to a common resolution, performing histogram equalization and applying a specific filter to remove small distortions without reducing the sharpness of the image.

4.2. Feature extraction

The main difficulty with any image retrieval operation is that the unit of information in image is the pixel and each pixel has properties of position and color value; however, by itself, the knowledge of the position and value of a particular pixel should generally convey all information related to the image contents [28]. To avoid this difficulty features are extracted using more than one way (i.e. HSV histogram, color coherence and color moments). This allows the system to extract from an image a set of numerical features, expressed as coded characteristics of the selected object, and used to differentiate one category from another.

4.3. Feature normalization

This step is used to prevent singular features from dominating the others and to obtain comparable value ranges. Here we normalize the features by a linear scaling procedure that provides for all the features to have values in a unit range. Let the lower and upper bounds for a feature component, x are a, b respectively. Then, the linear scaling is given as:

$$\varphi(x) = \frac{x-a}{b-a} \tag{6}$$

4.4. SNN image retrieval

In this step, the feature vectors, of the query image and an image in DB are fed into the SNN. Then the system predicts the corresponding metric vectors using a similarity function to retrieve the most relevant image according to a chosen threshold value. The similarity function between two images A and B can be defined as the weighted difference between their metric vectors as follows:

$$\mathcal{S}(A,B) = \varpi|\overrightarrow{v_A} - \overrightarrow{v_B}| \tag{7}$$

Furthermore, in this step, user can provide the system a feedback that updates the spline activation function and refine the unsatisfying retrieved results next times.

5. EXPERIMENTAL RESULTS

In this section, we demonstrate some experimental results to show the performance of the proposed SNNIR system.

5.1. Experiment 1

In this experiment, the designed cubic-splines network includes 69 inputs, 20 hidden units, and 5 outputs. In order to evaluate the learning performance of the cubic-splines network model, a comparison with standard sigmoidal NN was carried out. The runs consist in training splines network and sigmoidal network. Fig. 2 shows the plots of the Mean Squared Error (denoted as MSE) during the learning stage. The following formula of MSE was used:

$$MSE = \frac{1}{n}\left[\sum_{i=1}^{n}(y_i - d_i)^2\right]^{\frac{1}{2}} \tag{8}$$

where y_i, d_i and n are the actual network output, the desired output, and the number of samples respectively

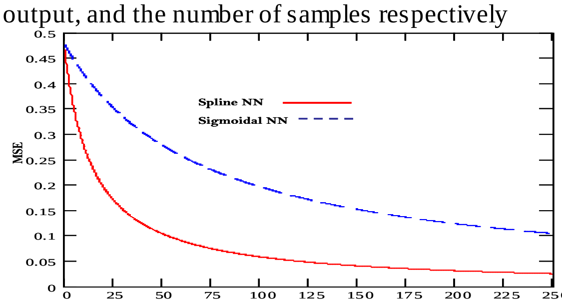

Fig. 2. Averaged learning curves comparison between cubic-splines NN and standard sigmoidal NN.

5.2. Experiment 2

To verify the performance of the proposed SNNIR system, the experiment results have been implemented with a general-purpose image DB including 500 images formed by five image classes: Building, Coast, Mountain, Car, and Grass. Each class depicts a distinct semantic topic. Some sample images have been selected to evaluate the efficiency of the proposed system. An image from each class was selected as a query image. In addition, the effectiveness of the system has been measured by both precision and recall. Table I depicts different retrieval results for each query image supported by the proposed system. The figures in table II show that the SNNIR does well in the terms of precision and recall in comparison to other CBIR systems like the SIMPLIcity system [29] and the RIBIR system [30]. Furthermore, the proposed system is very efficient for a set of images with same conception (see Fig. 3).

TABLE I

The results of retrieval for query images

Class	***Precision***	***Recall***
Building	0.89	0.96
Coast	1.00	1.00
Mountain	0.85	0.92
Car	0.90	0.72
Grass	1.00	0.92
Average performance	**0.93**	**0.90**

6. CONCLUSION

This paper has introduced the SNNIR system. By applying the splines network model, the system has determined nonlinear relationship between features so that more accurate similarity comparison between images was obtained. SNNIR system could also learn by feedback. Although the proposed retrieval approach has dealt with a five-class image database, it could be straightforwardly expanded to handle any larger database. Finally, the system could be easily extended to deal with video database.

ACKNOWLEDGMENT

This work is supported by BMBF Bernstein-Group (FKZ: 01GQ0702), EU grants (C4-NIMITEK 2, FKZ: UC4-3704M), Forschungspraemie (BMBF-Förderung, FKZ 03FPB00213) and Transregional Collaborative Research Centre SFB/TRR 62 ”Companion-Technology for Cognitive Technical Systems” funded by DFG.

REFERENCES

[1] S. Amari, “A universal theorem on learning curves,” *Neural Networks*, vol. 6, no. 2, pp. 161–166, 1993.

[2] M. Marchesi, G. Orlandi, F. Piazza, and A. Uncini, “Fast neuralnetworks without multipliers,” *IEEE Trans. Neural Networks*, vol. 3,pp. 101–120, Nov. 1992.

[3] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas imminent in nervous activity,” *Bull. Math. Biophys*, vol. 5, pp. 115–133, 1943.

[4] D. E. Rumelhart, J. L. McClelland, and the PDP Research Group, “Parallel Distributed Processing (PDP): *Exploration in the Microstructure of Cognition*,” vol. 1. Cambridge, MA: MIT Press, 1986.

[5] Widrow and M. Lehr, "30 years of adaptive neural networks: Perceptron, adaline and backpropagation," in *Proc. IEEE*, vol. 78, no. 9, pp. 1415-1442, 1990.

[6] C. Z. Tang and H. K. Kwan, "Convergence and generalization properties of multilayer feedforward neural networks," in *Proc. IEEE Int. Symp. Circuits Syst.*, San Diego, CA, 1992, pp. I-65–68.

[7] C. T. Chen and W. D. Chang, "A feedforward neural network with function shape autotuning," *Neural Networks*, vol. 9, no. 4, pp. 627–641, June 1996.

[8] G. Cybenko, "Approximation by superposition of a sigmoidal function," *Math. Contr. Signals Syst.*, no. 2, 1989.

[9] T. Chen and H. Chen, "Approximation of continuous functionals by neural networks with application to dynamic systems," *IEEE Trans. Neural Networks*, vol. 4, pp. 910–918, Nov. 1993.

[10] T. Chen, H. Chen, and R. Liu, "Approximation capability in C(Rn) by multilayer feedforward networks and related problems," *IEEE Trans. Neural Networks*, vol. 6, pp. 25–30, Jan. 1995.

[11] K. Hornik, M. Stinchombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, pp.359–366, 1989.

[12] Hornik, K., M. Stinchcombe, and H. White, "Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks," *Neural Networks*, vol. 3, pp. 551–560, 1990.

[13] M. Stinchcombe and H. White, "Universal approximation using feedforward networks with nonsigmoid hidden layer activation functions," in *Proc. IJCNN*, Washington, D.C., 1989, pp. I 613–617.

[14] Stinchcombe, M. and H. White, "Approximating and learning unknown mappings using multilayer feedforward networks with bounded weights," in *Proc. IJCNN*, San Diego, CA, 1990, pp. III 7–16.

[15] W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain , "Content-based image retrieval at the end of the early years," *IEEE Trans. on Pattern Recognition and Machine Learning*, vol. 22, no. 12, pp 1349-1380, 2000.

[16] X. Wan and C. C. J. Kuo, "A new approach to image retrieval with hierarchical color clustering," *IEEE Trans. on Circuits and Systems for Video Technology*, vol 8, no. 3, pp 628-643, 1998.

[17] AKulkarni, "CBIR Using Associative Memories," *Proceedings of the 6th WSEAS Int. Conference on Telecommunications and Informatics*, Dallas, Texas, USA, March 22-24, pp. 99-104, 2007.

[18] W. Hersh, H. Muller, *et al.* "Advancing biomedical image retrieval: development and analysis of a test collection," *J. Am. Med. Inform. Assoc.*, 13(5), pp. 488-496, 2006.

[19] Aisen AM, Broderick LS, *et al.*, "Automated storage and retrieval of thin-section CT images to assist diagnosis: System description and preliminary assessment," *Radiology*, 228, pp. 265-270, 2003.

[20] P. Schmid-Saugeon, J. Guillod, *et al.*, "Towards a computer-aided diagnosis system for pigmented skin lesions," *Computerized Medical Imaging and Graphics*, vol. 27, pp. 65-78, 2003.

[21] H. Müller, N. Michoux, D. Bandon, and A. Geissbuhler, "A review of content-based image retrieval systems in medicine – clinical benefits and future directions," *Int. J. Med. Inform.*, 73(1), pp.1-23, 2004.

[22] W. Hersh, J. Kalpathy-Cramer, *et al.* "Effectiveness of global features for automatic medical image classification and retrieval," *Pattern Recognition Letters*, vol. 29, pp. 2032-2038, 2008.

[23] Kalpathy-Cramer J, Hersh W, "Automatic Image Modality Based Classification and Annotation to Improve Medical Image Retrieval," *Proceedings of the 12th World Congress on Health (Medical) Informatics*, pp. 1334-1338, 2007.

[24] E. Catmull, R. Rom. "Computer Aided Geometric Design, chapter A Class of Local Interpolating Splines," *academic Press*, pp. 317–326, 1974.

[25] J. Hafner, H. S. Sawhney, W. Equitz, M.Flickner, and W. Niblack , "Efficient Color Histogram Indexing for Quadratic Form Distance Function, In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 17, No 7, pp. 729-739, July 1995.

[26] M. J. Swainand and D. H Ballard,. "Indexing Via Color Histograms,", *In Proceedings of third international conference on Computer Vision*, pp. 390-393, 1990.

[27] Ennesser, F., and Medioni G. "Finding Waldo, or Focus of Attention Using Local Color Information," *In IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 17, No 8, pp. 805-809, Aug 1995.

[28] R. Marmo *et al.* "Textural identification of carbonate rocks by image processing and neural network: Methodology proposal and examples," *Computers and Geosciences*, vol. 31, pp. 649–659, 2005.

[29] J.Z. Wang, J.Li, and G. Wiederhold, "SIMPLIcity: semantic-sensitive integrated matching for picture libraries, *IEEE Trans. Image Processing*, 9(1), pp.947-963, 2000.

[30] B. Si, W. Gao, H. Lu, W. Zeng, L. D, "An Image Retrieval Method Based Regions of Interest," *High Technology Letters*, vol.13, no.5, pp.13-18, 2003.

Fig. 3. Some retrieval results of SNNIR system.